3D-Model Search Engine from Photos *

Tarik Filali Ansary
LIFL UMR CNRS/USTL 8022
Univ. Lille 1 – France
filali@lifl.fr

Jean-Phillipe Vandeborre
LIFL UMR CNRS/USTL 8022
& GET/INT – Telecom Lille 1
Univ. Lille 1 – France
vandeborre@lifl.fr

Mohamed Daoudi
LIFL UMR CNRS/USTL 8022
& GET/INT – Telecom Lille 1
Univ. Lille 1 – France
daoudi@lifl.fr

ABSTRACT

In this paper, we present the FOX-MIIRE 3D-Model Search Engine. Our search engine is based on Adaptive Views Vlustering (AVC) algorithm [4]. The AVC method uses statistical model distribution scores to select the optimal number of views to characterise a 3D-model. The search engine also uses a probabilistic Bayesian method to retrieve 3D-models visualy similar to a query 3D-model, photos or sketches. We present our results on the *Princeton 3D Shape Benchmark database* (1814 3D-models). Our 3D-model search engine is available on-line to assess and test our results. To our knowledge the FOX-MIIRE search engine is the first search engine that accepts 3D-models retrieval from photos [5]. The FOX-MIIRE search engine adapts its own user interface depending on the web access device (desktop computer, PDA, SmartPhone).

Categories and Subject Descriptors

H.3 [**Information Systems**]: Information Storage and Retrieval; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing—*Indexing methods*; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval; H.3.5 [**Information Storage and Retrieval**]: Online Information Services—*Web-based services*

Keywords

3D-model indexing, view-based approach, Bayesian indexing, 3D-search engine, PDA, SmartPhone.

1. INTRODUCTION

The development of 3D modeling and digitalising technologies has made the 3D-model generation process much easier. Moreover, thanks to the Internet, users can download a large number of free 3D-models from all over the world. This has increased the need for developing efficient techniques for content-based 3D-model retrieval.

Recently, some experimental 3D-shape search engines have been made, such as the 3D-model search engine at the *Princeton University* [6], the 3D-model retrieval system at the *National Taiwan University* [3], the Ogden IV system at the *Mational Institute of Multimedia Education*, Japan, and the 3D-model similarity search engine at the *University of Konstanze*.

To search a database for 3D-models that are visually similar to a view, a sketch or a photo of a 3D-model is a very intuitive way. But, it is a challenging problem. The main idea in 3D retrieval using 2D-views or photos is that two 3D-models are similar if they also look similar from different angles. So, the proposed solutions are to match one or more photos (or sketches or views) to the 3D-models they are similar to.

However, to our knowledge no on-line 3D-model search engine can retrieve 3D-models from one or more photos. A complete survey on 3D shape retrieval can be found in Tangelder and Veltkamp [9].

In this paper, we propose a method for 3D-model retrieval from one or more photos (photographs, sketches, views) based on 2D-views. The method aims at providing an optimal selection of 2D-views from a 3D-model, and a probabilistic Bayesian method for 3D-model indexing from these views. This paper is organised as follows. In section 2, we present the main principles of our method for characteristic view selection. In section 3, our probabilistic 3D-model retrieval from photos is presented. Then, the results obtained on a database containing 1814 3D-models (*Princeton 3D Shape Benchmark database*) are discussed demonstrating the performance of our method. Finally, we present our on-line 3D search engine and three usage scenarios.

2. SELECTION OF CHARACTERISTIC VIEWS

Let $D_b = \{M_1, M_2, \ldots, M_N\}$ be a collection of N three-dimensional models. We want to represent each 3D-model M_i by a set of 2D-views that best represents it. To achieve this goal, we first generate an initial set of views from the 3D-model, then we reduce this set to only those that best characterise this 3D-model. In this section, we present our algorithm for characteristic view selection from a three-dimensional model.

2.1 Generating the initial set of views

To generate the initial set of views for a model M_i of the collection, we create 2D-views (projections) from multiple viewpoints. These viewpoints are equally spaced on the unit sphere. In our current implementation, we use 320 initial views. The views are silhouettes only, which enhances the efficiency and the robustness of the image metric. To

*This work has been partially supported by the European Network of Excellence Delos No. 507618 http://www.delos.info.

CIVR'07, July 9–11, 2007, Amsterdam, The Netherlands.
ACM 978-1-59593-733-9/07/0007.

represent each of these 2D-views, we use 49 coefficients of Zernike moment descriptor [7]. Due to the use of Zernike moments, the approach is robust against translation, rotation, and scaling.

2.2 Characteristic view selection

As every 2D-view is represented by 49 Zernike moment coefficients, choosing a set of characteristic views that best characterise the 3D-models (320 views) is equivalent to choose a subset of points that represents a set of 320 points in a 49-dimension space. Choosing X characteristic views which best represent a set of $N = 320$ views is well known as a *clustering problem*.

One of the widely used algorithm in clustering is *K-means* algorithm. Its attractiveness lies in its simplicity and in its local-minimum convergence properties. However, it has one main shortcoming: the number of clusters K has to be supplied by the user.

As we want from our method to adapt the number of characteristic views to the geometrical complexity of the 3D-model, we use a method derived from K-Means called *X-Means* [8]. Instead of using a fixed number of clusters, we propose to use a range in which we will choose the "optimal" number of clusters. In our case, the range will be $[1, \ldots, 40]$. In our system, we selected 40 as the maximum number of characteristic views. This number of views is a good compromise between speed, descriptor size and representation.

To select the best characteristic view set, we use the *Bayesian Information Criteria* (BIC), which scores how likely the representation model (using one, two, or Xcharacteristic views) fits the data. The best characteristic view set will be the one that gets the highest BIC score. Details of the *Adaptive Views Clusterings* method can be found in [4, 5].

3. PROBABILISTIC APPROACH FOR 3D INDEXING

The main idea of our probabilistic approach is that **not** all views of a 3D-model have the same importance. There are views which represent the 3D-model better than others. On the other hand, simple objects (e.g. cube, sphere) can be at the root of more complex objects, so they have a higher probability to be relevant. In this section, we present a probabilistic approach that takes into account that views do not have the same importance, and that simple objects have higher probability to appear than more complex one.

Each model of the collection D_b is represented by a set of characteristic views $V = \{V_1, V_2, \ldots, V_C\}$, with C the number of characteristic views. To each characteristic view corresponds a set of represented views called V_r.

As mentioned before, we want to find the 3D-models that correspond to one or more request photos or views from a 3D-model. We assume that in a query $Q = \{I_1, I_2, \ldots, I_K\}$ all the K images represent the same object.

Considering a query Q, we wish to find the model $M_i \in D_b$ which is the closest to the query Q. This model is the one that has the highest probability $P(M_i|Q)$. Knowing that a query is composed of one or more images or a 3D-model views, the probability $P(M_i|Q)$ can be written as follows:

$$P(M_i|Q) = \sum_{k=1}^{K} P(V_Q^k|Q)$$

$$Max_j\left(\frac{P(h_j^k|V_{M_i}^j, M_i)P(V_{M_i}^j|M_i)P(M_i)}{\sum_{i=1}^{N}\sum_{j=1}^{\hat{v}} P(h_j^k|V_{M_i}^j, M_i)P(V_{M_i}^j|M_i)P(M_i)}\right).$$

With $P(M_i)$ the probability to observe the 3D-model M_i, $P(V_{M_i}^j|M_i)$ the probability to observe the characteristic view j of the model M_i. The value $P(h_j^k|V_{M_i}^j, M_i)$ is the probability that, knowing that we observe the characteristic view j of the model M_i, this view corresponds to image k of the request Q

To summarise, in this section we presented our Bayesian retrieval framework which takes into account the number of characteristic views of the 3D-model and the importance (amount of information) of its views. More details about the Bayesian approach can be found in [4].

4. EXPERIMENTAL RESULTS

In this section, we present the experimental process and the obtained results. The method we described in the previous sections have been implemented using C++. The system consists of an off-line characteristic view extraction algorithm and an on-line retrieval process.

In the off-line process, the characteristic view selection takes about 18 seconds per model on a PC with a Pentium IV 2.4 GHZ CPU. In the on-line process, the comparison takes less than 1 second for the entire collection of 1814 3D-models. Figure 1 shows the FOX-MIIRE 3D-model search engine architecture.

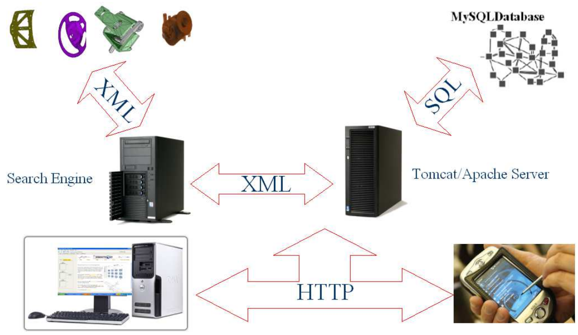


Figure 1: FOX-MIIRE 3D-model search engine architecture.

To evaluate our method, we used the *Princeton Shape Benchmark database* [2], a standard shape benchmark widely used by the shape retrieval community. The *Princeton Shape Benchmark database* appeared in 2004 and is one of the most exhaustive benchmarks for 3D-shape retrieval. It contains 1814 classified 3D-models collected from 293 different web domains. There are many classifications given to the objects in the database. Most classes contain objects with a particular function (e.g cars). Yet, there are also cases where objects with the same function are partitioned in different classes based on their shapes (e.g, round tables versus rectangular tables).

Based on some standard measures, we compared our method to thirteen state-of-the-art methods on *Princeton 3D Shape*

Benchmark database. Our method gives the second-best results. Our AVC method gives a good quality/cost compromise compared to other well-known methods. The results of our method on a large 3D-CAD-model database (5000 models) supplied by the car manufacturer *Renault*, show that our method can also be suitable for 3D-CAD-model retrieval. Our method is robust against noise and model degeneracy. It is usable in the case of topologically ill-defined 3D-model. A practical 3D-model retrieval system based on our approach is available on the web for on-line tests [1].

The experiments on *Princeton Shape Benchmark Database* – composed of 1814 3D-models – show the good retrieval results using one or more photos. The precision gain from using more than one photo can be up to 78% [5].

5. USAGE SCENARIOS

To experiment our algorithms and to assess the results presented in the previous sections, we developed an on-line 3D search engine. Our search engine can be reached from any device having compatible web browser (desktop computer, PDA, SmartPhone, etc.) [1]. The FOX-MIIRE 3D Search Engine offers three usage scenarios:

1. **Searching 3D-model database using a 3D-model query.** The FOX-MIIRE 3D-search engine offers to browse the 3D-model database using the database classification. The user can retrieve the 3D-models that match a selected 3D-model using a click on the correspond search button (Figure 2).

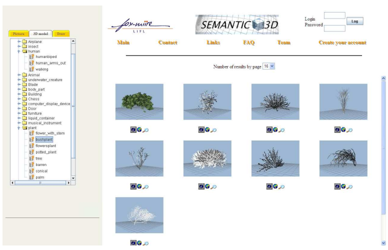


Figure 2: Browsing the 3D-Models database using the database classification.

2. **Searching 3D-model database using a sketch query.** Using the drawing interface, the user can draw one or more sketches representing the 3D-model and using the search button the user can retrieve the 3D-models corresponding to these sketches. Figure 3 shows a sketch of a *dog* and the corresponding 3D-models retrieved from the 3D-model database. Figure 4 shows the results of the 3D-models retrieved using the fifth 3D-model retrieved (in Figure 3) as a query.

3. **Searching 3D-model database using a photo query.** To our knowledge, the FOX-MIIRE search engine is the world first search engine to offer 3D retrieval from one or more photos. The user can retrieve the 3D-models that match a photo loaded from his/her personal computer or from the photo query historic available in the web interface. Figure 5 shows a photo of a *bike* and the corresponding 3D-models retrieved from the 3D-model Database. Figure 6 shows the results of using two photos. Using more photos give a more accurate result as detailled in [5]. The gain from using multiple images in a query faces the problem of "how to get these query images?" Using web images search-engines, camera or hand-drawing can solve the problem, but it is still time and effort consuming to get or draw five or more images. We believe that using two or three photos as a maximum makes a good compromise between time-effort and accuracy [5].

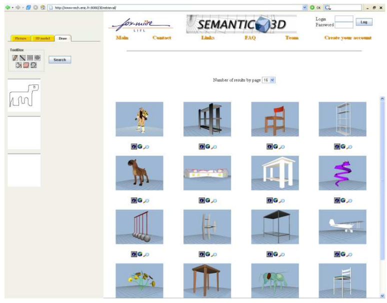


Figure 3: 3D retrieval results using a sketch.

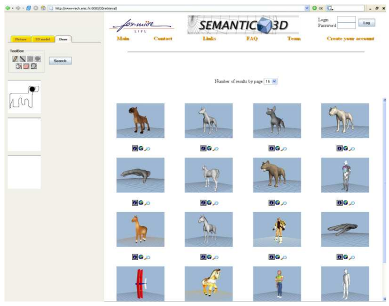


Figure 4: Browsing the 3D-model database using a 3D-model as a query.

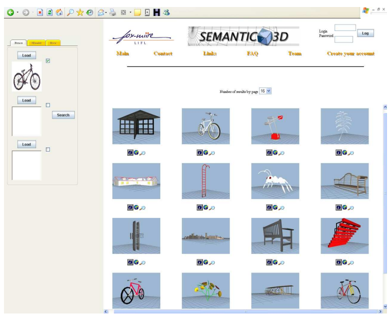


Figure 5: 3D retrieval results using one photo.

Our work can be applied to CAD-model retrieval from photos and can be adapted to e-shopping, where instead of browsing big catalogues of products, the user presents one or more photos of a similar object and the search engine will return the most relevant results.

Depending on the web access device he/she is using, the user faces two different kinds of web interfaces: a rich web interface for full-featured web-browsers, and a simpler interface for mobile device (PDA, SmartPhone) web-browsers. In both cases, the results returned by the 3D search engine

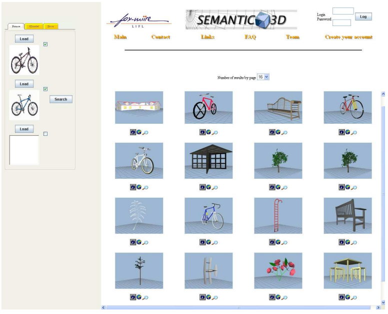


Figure 6: 3D retrieval results using two photos.

are the same. The only difference lies in the design of the result presentation.

Figure 7(a) shows the user interface where the he/she can upload photos to start the 3D-model retrieval. Figure 7(b) shows the results of the query in figure 7(a). Figure 8(a) shows some details of a 3D-model selected by the user. As most of the PDA/SmartPhone cannot display 3D-models, Figure 8(b) shows the interface we developped to “turn” around the 3D-model using pre-calculated images.

6. CONCLUSION

In this paper, we presented a 3D-model search engine from 3D-models, sketches and photos as queries. The search engine is based on a characteristic view selection algorithm – called *Adaptative Views Clustering* – that relates the number of views to the geometrical complexity of the 3D-models. The number of characteristic views varies from 1 to 40 per model. We also use a new probabilistic retrieval approach that corresponds one or more photos, representing the same object, to a 3D-model characteristic views. The experiments of our method on the *Princeton Shape Benchmark Database* (1814 3D-models) show the good retrieval results using one or more photos. The FOX-MIIRE 3D-model retrieval search engine is available on the Web for on-line tests [1]. To our knowledge, it is the first 3D-model retrieval system from photos. It is also the first 3D-model retrieval search engine for mobile devices such as PDA or SmartPhones.

7. REFERENCES

[1] FOX-MIIRE 3D retrieval system. http://www-rech.telecom-lille1.eu/3dretrieval.

[2] Princeton Shape Benchmark, 2004. http://shape.cs.princeton.edu/benchmark.

[3] D. Y. Chen, X. P. Tian, Y. T. Shen, and M. Ouhyoung. On visual similarity based 3D model retrieval. In *Eurographics*, volume 22, pages 223–232, 2003.

[4] T. Filali Ansary, M. Daoudi, and J.-P. Vandeborre. A bayesian 3D search engine using adaptive views clustering. *IEEE Transactions on Multimedia*, 9(1):78–88, January 2007.

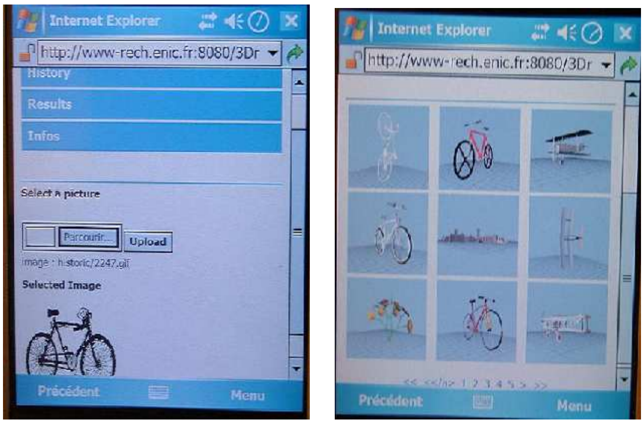
[5] T. Filali Ansary, J.-P. Vandeborre, and M. Daoudi. On 3D retrieval from photos. In *3rd IEEE International Symposium on 3D Data Processing Visualization Transmission (3DPVT'06)*, Chapel Hill, North Carolina, USA, June 14-16 2006.

[6] T. Funkhouser, P. Min, M. Kazhdan, A. Haderman, D. Dobkin, and D. Jacobs. A search engine for 3D models. *ACM Transactions on Graphics*, 22(1):83–105, 2003.

[7] A. Khotanzad and Y. H. Hong. Invariant image recognition by Zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):489–497, May 1990.

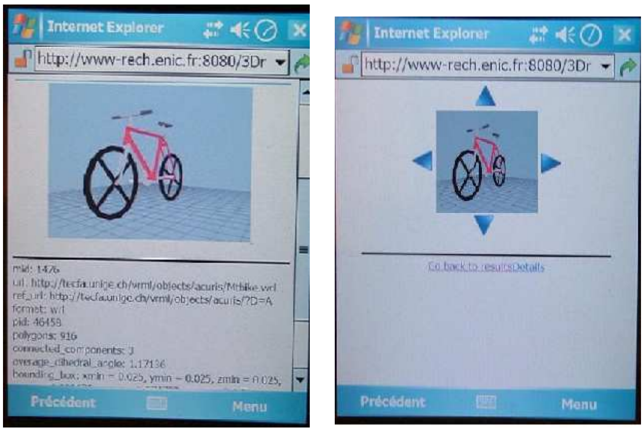
[8] D. Pelleg and A. Moore. X-means: Extending k-means with efficient estimation of the number of clusters. In *International Conference on Machine Learning*, pages 727–734, 2000.

[9] J. W. H. Tangelder and R. C. Veltkamp. A survey of content-based 3D shape retrieval methods. In *IEEE Shape Modeling International*, pages 145–156, 2004.



(a) (b)

Figure 7: (a) The interface to upload and search for 3D-models from a mobile device. (b) 3D-models retrieved from the *bike* photo.



(a) (b)

Figure 8: (a) Details of the selected 3D-model. (b) The "Turn Around" intereface for mobile device.